

**High Mobility Heterojunction Complementary Field Effect Transistors and  
Methods Thereof**

**FIELD OF THE INVENTION**

5 The present invention relates to electronic devices and systems. In particular it relates to field effect devices, and a method of producing such structures, with a heterojunction source/drain which substantially coincides with the metallurgical junction between the source/drain and the device body.

**BACKGROUND OF THE INVENTION**

10 Today's integrated circuits include a vast number of devices. Smaller devices are key to enhance performance and to improve reliability. As MOS (Metal Oxide Semiconductor Field- Effect- Device, a name with historic connotations meaning in general an insulated gate Field- Effect- Device) devices are being scaled down, the technology becomes more complex and new methods are needed to maintain the expected performance enhancement from one generation of devices to the next.

15 Silicon MOS scaling has become a major challenge in the semiconductor industry. Traditional techniques start to fail in reducing certain undesirable physical effects as device dimensions shrink down to the nanometer regime. For example,

anti-punchthrough or halo implantation is used to reduce the short-channel effects.

However, the abrupt doping profiles are difficult to achieve due to temperature enhanced diffusion, and these highly doped channels or pocket implant regions not only increase junction capacitance and band-to-band tunneling, but also degrades carrier mobility in the channel.

Band-gap engineering, namely introducing new materials into Si processing with different energy bands than Si, gives an important additional degree of freedom in device design. Among such new materials the SiGe alloy is one of the prominent members. The growth of high-quality compressively strained SiGe material by molecular beam epitaxy (MBE), or various types of chemical vapor deposition (CVD), allows incorporation of band-gap engineering concepts into a mature silicon technology.

A novel way to reduce short-channel effects is to have a built-in energy barrier at the source/body junction. As the height of the heterojunction barrier does not depend on the applied bias, it can resist the drain induced barrier lowering (DIBL). The band offset provided by SiGe heterojunction is mostly in the valence band, and it is very suitable for using such an effect for PFETs. (In the following the terms PFET and PMOS, as well as the terms of NFET and NMOS will used interchangeably.)

Heterojunction MOSFETs (HJMOSFETs) have been disclosed, for example, in US. Patent 6,319,799 B1 "High Mobility Heterojunction Transistor and Method" by Q. Ouyang, et al. and in a simulation study in "A Novel Si/SiGe Heterojunction pMOSFET with Reduced Short-Channel Effects and Enhanced Drive Current," IEEE Transactions on

Electron Devices, V. 47, p. 1943 (2000), by Q. Ouyang, et al. This latter reference showed that the SiGe/Si heterojunction and the dopants metallurgical junction have to coincide with each other to a relatively high precision, or the p-dopant has to be contained within the SiGe regions, in order to maintain the valance band offset between the source and body. Only when such precision is achieved can the heterojunction be effectively used to reduce the off state leakage and short-channel effects. Hitherto there was no method to give the desired coincidence of the heterojunction and the metallurgical junction, and no lateral device structure with such coincidence has been disclosed.

### SUMMARY OF THE INVENTION

The present invention describes heterojunction source/drain MOSFET devices in which the PMOS is a buried SiGe channel device, while the NMOS is a surface silicon channel device. According to one aspect of the invention, the heterojunction and metallurgical junction of the device substantially coincide to within a close tolerance. Another aspect of the invention relates to a method comprising forming heterojunction source/drain by epitaxial deposition. The new heterojunction source/drain MOSFET devices can be realized in silicon on insulator (SOI), or in bulk Si technologies. The SiGe epitaxial source/drain will be compressively-strained due to lattice constant mismatch between Si and SiGe. Also, since the hole current in the PMOS will be primarily confined to a buried compressively strained SiGe channel, due to the high hole mobility in such a

materials, this PMOS device can have potentially similar current carrying capability as the NMOS device. The high quality PMOS leads to excellent CMOS circuit capabilities for building processors.

5 In accordance with the objectives listed above, the present invention describes Si based MOSFET devices with a SiGe heterojunction source/drain where the heterojunction and the metallurgical junction of the source/drain and Si body are coinciding with a minimal tolerance.

10 It is a further object of the present invention to teach a method for any field effect device for processing a heterojunction source and drain with a minimal tolerance in the heterojunction and the metallurgical junction lineup.

It is a further object of the present invention to teach processors which comprise chips containing such field effect devices having heterojunction source and drain with a minimal tolerance in the heterojunction and the metallurgical junction lineup.

#### BRIEF DESCRIPTION OF THE DRAWINGS

15 These and other features of the present invention will become apparent from the accompanying detailed description and drawings, wherein:

Fig. 1 schematically shows exemplary embodiments of a Si based heterojunction source/drain field device;

Fig. 2 schematically shows exemplary embodiments of a Si based heterojunction

source/drain field device in complementary configurations;

Fig. 3 shows processing steps in the fabrication of an exemplary embodiment of a heterojunction source/drain field device;

Fig. 4 shows processing steps in the fabrication of an alternate exemplary embodiment of a heterojunction source/drain field device; and

Fig. 5 shows a symbolic view of a processor containing a chip which contains a heterojunction source/drain field device.

### DETAILED DESCRIPTION OF THE INVENTION

Fig. 1 schematically shows exemplary embodiments of a Si based heterojunction source/drain field device. Each embodiment contains the feature of the source and drain of the field effect device (FET) forming a heterojunction with the body of the device. As every FET, the device has (at least one) gate 52. The conductive gate 52 in state of the art technologies is protected by a cover layer 51, which in Si based technologies is typically oxide and SiN. The gate 52 is electrically isolated from the rest of the device by the gate dielectric 53. The common property of FET-s is that the gate 52, by capacitive coupling across the gate insulator 53, controls the flow of the device current between the source and drain. Insulator regions 54 serve for isolating the device. Typically the isolation is accomplished by shallow trenches 54, usually made of SiO<sub>2</sub>, as shown in the figures.

The material of the structure shown on Fig. 1 belongs in the broad category of Si based materials. In microelectronics technology the material that has progressed the farthest in miniaturization is silicon (Si). Si based materials are various alloys of Si in the same basic technological content as Si. One such Si based material of importance for microelectronics is the silicon-germanium (SiGe) alloy. SiGe has a larger lattice constant than Si, increasingly so with increasing Ge concentration. Accordingly, when SiGe is epitaxially grown or deposited on Si it is under compressive strain. The bandgap of SiGe is smaller than that of Si. The higher the Ge content the smaller the SiGe bandgap. At a Si-SiGe heterostructure the bandgap difference is almost exclusively accommodated in the valence band. In the conduction band the band-discontinuity is almost negligible. In an FET having a source and a drain in a heterojunction relationship with the device body has advantages for miniaturization of the device. Also, having a channel consisting of strained SiGe material is very advantageous for carrier mobility, especially for holes. The reasons for such advantages and their details can be found in publications: "A Novel Si/SiGe Heterojunction pMOSFET with Reduced Short-Channel Effects and Enhanced Drive Current," IEEE Transactions on Electron Devices, V. 47, p. 1943 (2000), by Q. Ouyang, et al. Also, a PMOS with SiGe heterojunction source/drain and SiGe channel was disclosed in the US Patent 6,319,799 B1 "High Mobility Heterojunction Transistor and Method" to Q. Ouyang, et al., incorporated herein by reference.

In the representative embodiments shown on Fig. 1 the device has a crystalline body 40 essentially consisting of Si. A SiGe layer 20 is epitaxially disposed on the body

40. The terms “epitaxially”, “epitaxy”, “epi” etc. carry their customary usage: meaning the single crystal lattice structure carries across an interface. Typically a single crystal material forms a platform onto which another single crystal material, with matching crystalline characteristics, is deposited by one of several techniques known in the art. Such techniques are, for instance, molecular beam epitaxy (MBE), or various types of chemical vapor depositions (CVD). Because of the lattice constant relationship between the SiGe layer 20 and the Si body 40 the SiGe layer 20 is under compressive strain. Such a strain enhances significantly the hole mobility. On top of the SiGe layer 20 a Si layer 30 is epitaxially disposed. This Si layer 30, for example, can serve one or more purposes such as: 1) as a channel for electron type device current; 2) allows for growth of good quality gate dielectric 53, typically of SiO<sub>2</sub>, possibly in a mixture with SiN; or 3) as a protective layer to minimize contamination of the SiGe layer 20.

The border surface between the crystalline Si body 40 and the SiGe source and drain 10, which are in an epitaxial relation with the Si body, form a heterojunction interface 11. The term heterojunction means that two dissimilar materials meet and the energy bands have abrupt discontinuity at that interface 11. There is another interface of importance in the device, which is called a metallurgical junction. This is the surface, shown with a dashed lined 65 in the embodiments of Fig. 1., where the conductivity type changes between the body 40 and the source and drain 10. The body 40 and the source/drain have opposite conductivity type. If the source/drain 10 is p-type then the body is n-type, and vice versa, if the source/drain 10 is n-type then the body is p-type.

Where the two type of dopants, the p-type, such as boron (B), and n-type, such as phosphorous (P), or arsenic (As), balance each other out, that is the location of the metallurgical junction 65 between the body region 40 and the source/drain region 10. In order for the potential advantages of a heterojunction source/drain to materialize, it is preferable that the heterojunction 11 substantially coincides with the metallurgical junction. Thus, one aspect of the present invention provides a method of fabricating the representative FET device embodiments of Fig. 1 with the heterojunction 11 and the metallurgical junction 65 coinciding with a tolerance of less than about 10nm, and preferably less than about 5nm. Although in Figs. 1A, 1B, and 1C the metallurgical junction is shown on the body 40 side of the heterojunction 11, in reality it can be on the source/drain10 side of the heterojunction 11 as well. The main point is that the two are coinciding with a minimal tolerance.

In all representative embodiments presented, the Ge concentration in the source and drain 10 and the SiGe epitaxial channel layer 20 is in the 15% to 50% range, preferably between about 20% to 40%. The SiGe thicknesses in the source/drain 10 are staying below the critical thickness for the given Ge concentration. The critical thickness is defined such that above it the SiGe will relax and defects and dislocations will form. The thickness of the SiGe epitaxial layer 20 typically is between about 5nm and 15nm. The thickness of the epitaxial Si layer 30 is typically between about 5nm and 15nm. Fig. 1A shows an embodiment where the body is bulk Si. These type of devices are the most common devices in present day microelectronics. Figs. 1B and 1C show representative



embodiment of the heterojunction source/drain FET device when the Si body 40 is disposed on top of an insulating material 55. This type of technology is commonly referred to as silicon on insulator (SOI) technology. The insulator material 55 usually, and preferably, is SiO<sub>2</sub>. Fig. 1B shows an SOI embodiment where the body 40 has enough  
5 volume to contain mobile charges. Such SOI devices are called partially depleted devices. Fig. 1C shows an SOI embodiment where the volume of the body 40 is insufficient to contain mobile charges. Such SOI devices are called fully depleted devices. For devices shown in Fig. 1B and 1C there is, at least a thin, layer of body underneath the source and drain 10. This body material serves as the seed material onto which the epitaxial SiGe  
10 source and drain 10 are grown. In an alternate embodiment, for extremely thin fully depleted SOI devices, one could grow the source and drain 10 laterally, from a lateral seeding, in which case the source and drain 10 would penetrate all the way down to the insulating layer 55.

Fig. 1A shows an additional feature of representative embodiments a Si cap layer  
15 15 on top of the SiGe source and drain 10. The purpose of such a Si cap layer 15 is mainly to facilitate better electrical contact to the source and drain 10. The thickness of the Si cap layer can be relatively relaxed, typically between about 2nm and about 30nm. This Si cap layer 15 is typically formed epitaxially on the SiGe source and drain 10, immediately following the SiGe epitaxial growth. Although the Si cap layer 15 is  
20 indicated only in Fig. 1A, somebody skilled in the art would understand, that it can be part of the other representative embodiments, as well.

Since the fully depleted SOI devices as in Fig. 1C have a shallow profile, it can be advantageous to have a so called raised source/drain 10 in such structures for reducing source/drain series resistance. FET devices have a well defined surface plane, shown by the dashed line 60. This top surface plane approximately passes across the interface  
5 between the gate dielectric 53 and silicon layer 30. For the fully depleted SOI FET of Fig. 1C the source10 and drain10 are raising above the top surface plane, yielding the desired advantages.

In the representative embodiments shown in Fig. 1, the Si body 40 can have n-type conductivity, namely the device is a PFET, and in these devices a hole device  
10 current flow, as a consequence of the heterojunction discontinuity in the valence band, is confined predominantly in the SiGe layer 20. Since a heterojunction 11 is desirable between a body 40 and a source/drain 10, but it could be disadvantageous if it formed a barrier in the path of the device current, it is desirable for the SiGe layer 20 to substantially equal the Ge concentration of the SiGe in the source 10 and the drain 10.  
15 With substantially equal Ge concentrations, the holes can pass between the source 10 and drain10 an the SiGe layer 20 without hindrance.

Alternatively, the Si body 40 can have p-type conductivity, namely the device is an NFET. For these devices, the electron device current flow, as a consequence of the lack of heterojunction discontinuity in the conduction band, is confined predominantly in  
20 the Si layer 30.

Fig. 2 schematically shows exemplary embodiments of a Si based heterojunction

source/drain field device in complementary configurations. Fig. 2A shows two heterojunction source/drain field devices in complementary configuration (CMOS), namely p-type MOS and n-type MOS. One can have a CMOS where both type of devices are of the same character - i.e., both PMOS and NMOS having heterojunction source/drains , with one device having a n-type conductivity Si body 40, and the second device having a p-type conductivity Si body 40'. In one embodiment, both the PMOS and the NMOS have the heterojunction and the metallurgical junction coinciding with a tolerance of less than about 10nm, and preferably less than about 5nm. Alternatively, one can have CMOS configurations in which both PMOS and NMOS have SiGe heterojunction source/drain, but only the PMOS device have the substantial coincidence, less than about 10nm, and preferably less than about 5nm of tolerance between the heterojunction and the metallurgical junction.

Fig. 2B shows another representative embodiment when only the PMOS device has a heterojunction source/drain 10 with the substantially coinciding heterojunction and metallurgical junction within less than about 10nm, and preferably less than about 5nm of tolerance. In this case, however, the NMOS device does not have SiGe in the source 70 and drain 70, although it has the SiGe layer 20 and Si layer 30. Omitting the SiGe source/drain 10 from the NMOS device may have advantages. For example, the compressively strained epitaxial SiGe source/drain will cause some compressive stress in the silicon body near the gate edges after the cooling down from the epitaxial SiGe growth temperature, typically about at 500-600°C. This will cause the SiGe near the gate

edges to be even more compressively strained. This additional compressive strain is actually advantageous for the PMOS, resulting in even higher hole mobility in the p-channel. However, the compressive strain in the Si n-channel of the NMOS will lower the electron mobility compared to bulk silicon. The effect may become more dominant for shorter gate lengths. Therefore, this representative embodiment of a CMOS in Fig.2B where the NMOS has conventional source and drain 70 with the SiGe layer 20 and Si layer 30, has superior performance due to the strength of the PMOS device with a heterojunction source/drain 10. In a further representative embodiment, the PMOS device with a heterojunction source/drain 10 and the substantially coinciding heterojunction and metallurgical junction to within less than about 10nm, and preferably less than about 5nm of tolerance, can be paired up with any conventional NMOS device to form a CMOS configuration.

One advantage of having a PFET device with a heterojunction source/drain and with the substantially coinciding heterojunction and metallurgical junction is that additionally to the built-in potential by the p/n junction between the n type Si body 40 and the p-type SiGe source/drain 10, the band offset at the strained SiGe/Si interface 11 provides a potential barrier for holes and it does not vary with the drain bias such that the drain-induced barrier lowering and off-state leakage current can be reduced significantly for the PFET.

The advantage to the PFET arising from the increased hole mobility can be further enhanced by orienting devices according to crystallographic planes and directions. Since

hole mobility is known to be generally higher on a (110) surface and electron mobility is known to be generally higher on a (100) surface, the CMOS may have a hybrid crystal orientation configuration such that the PFET is on (110) surface and the NFET is on (100) surface. In general it can be advantageous to orient both the PMOS and NMOS devices with their surface plane 60 (Fig. 1C) lying essentially in any of the main crystallographic directions: (100), (110), and (111). In addition, on a typical (100) wafer surface, the channels can be laid out such that the current flow is along the  $\langle 100 \rangle$  and/or  $\langle 110 \rangle$  direction, whichever has the higher carrier mobility for holes and electrons. The local stress may have a dependency on the channel direction and crystal orientation.

The advantages for both PFET and NFET devices having a strained SiGe source/drain are numerous, as recounted now. The part of device series resistance which is caused by the source/drain is lower due to higher solubility of B and P in SiGe compared to Si. Next, the contact resistance can be smaller due to the smaller band gap of SiGe. Furthermore in SOI embodiments floating body effects are reduced because the hot carriers generated by impact ionization near the drain can diffuse through the barrier into the source due to the reduced barrier height of the narrower band-gap in SiGe. This effect occurs in both PFET and NFETs. Also, the source/drain breakdown (BVds) can significantly be improved. For ultra-thin fully depleted SOI, the raised SiGe source/drain of Fig. 1C can be used. The strain in SiGe will be maintained as long as its thickness is under the critical thickness.

One aspect of the present invention provides for the fabrication of the

heterojunction source/drain for FETs by epitaxial deposition of the source/drain material. Such epitaxial deposition can be achieved with desired material purity, crystalline quality, and control, for instance, by molecular beam epitaxy (MBE), or various types of chemical vapor deposition (CVD).

5               Such a method of fabricating source/drain by epitaxial deposition is not confined to the Si, SiGe material system, but can be applied to a wide variety of devices, for instance for the III-V semiconductor family. Thus, a source and a drain are formed generally by epitaxial deposition of a first material which forms a heterojunction with a second material, which second material constitutes the body of the FET device. If  
10           advantageous, such as in the case for the SiGe source/drain PMOS of Fig. 1, one can further include the fabrication step of providing a channel between the source and the drain, such that the channel consists essentially of the first material, namely the material of the source/drain. Alternatively, or together with the first material channel, one can carry out the step of providing a channel between the source and the drain which is  
15           essentially of the second material, namely the same as the device body. Furthermore, the precise control afforded in MBE and/or CVD allows for doping the source and drain such that the body is provided with one conductivity type and the source and the drain are provided with an opposing conductivity type, forming a metallurgical junction between the body and the source and the drain, and furthermore having the heterojunction and the  
20           metallurgical junction coinciding with a close tolerance. Such a close tolerance in state of the art FETs is desirable to be less than about 10nm, and preferably less than about 5nm.

Fig. 3 and Fig. 4 show processing steps in the fabrication of exemplary embodiments of a heterojunction source/drain field device. The CMOS processing of a SiGe heterojunction source/drain field device can follow the many well established steps of CMOS processing known in the art. Only those steps that are unique to the heterojunction source/drain device are discussed and shown on Fig. 3 and Fig. 4. Fig. 3 and Fig. 4 show processing steps for the exemplary embodiment of a non-fully depleted SOI devices, as on Fig. 1B. However, in other exemplary embodiments, such as bulk devices, or fully depleted SOI devices, the specific steps for the heterojunction source/drain are the same.

Fig. 3A shows the fabrication after performing shallow trench isolation in the Si body 40 which is disposed on an insulating material 55, and having the SiGe epi layer 20 and Si epi layer 30 already deposited. The order of isolation and epi growth can be selected to suit a particular desired processing sequence. One can carry out the growth of the epi layers 20 and 30 before, or after, the fabrication of the shallow trench isolation 54.

Fig. 3B shows the process after several more processing steps known in the art have been carried out. CMOS well implants and threshold adjust implants have been done, accordingly the Si body becomes of two type of conductivity n and p, type 40 and 40'. Gate dielectric 53, typically an oxide, has been grown. Gates 52 are deposited and patterned, and gate covers/spacers 51 are formed. Fig. 3C shows the source/drain regions being recessed to make room for the subsequent deposition of SiGe material. Such recessing/etching can be carried with techniques known in the art, such as reactive ion

etching, and/or various wet etchings.

Fig. 3D shows a selective epitaxy with in-situ p-doped SiGe alloy into the recessed source/drain regions. The p- doped epi goes into both the PMOS and the NMOS yielding the identical source/drains 10 in both devices. The in-situ doped deposition of the SiGe alloy allows for sufficient control of the doping profiles resulting in substantial coincidence between the heterojunction and the metallurgical junction. In an exemplary embodiment this selective epitaxy of the source and drain can be carried out by ultra high vacuum chemical vapor deposition (UHV- CVD) with  $\text{SiH}_4$ ,  $\text{Si}_2\text{H}_4$ , or  $\text{GeH}_4$  precursor in a temperature range of about 400-650°C, typically about 550°C. To control the metallurgical junction position, the film can be in-situ boron doped with  $\text{B}_2\text{H}_6$  added during the epitaxy.

This figure also indicates representatively a capping step yielding epitaxial Si cap layers 15 on top of the SiGe source and drain 10. The epitaxy of the Si cap layer 1 follows the epitaxy of the SiGe epitaxy of the source and drain. Such Si cap layers can facilitate contact to the device. The thickness of the Si cap layer can be relatively relaxed, typically between 2nm and 30nm. Although the step of depositing such a Si cap layer 15 is indicated only in Fig. 3D, somebody skilled in the art would notice, that it can part of the fabrication sequences of other representative embodiments, as well.

Fig. 3E shows the final steps in such a preferred embodiment. While the PMOS is masked 57, the NMOS source/drain regions are implanted 58 with an n-type species, such as phosphorus, or arsenic for example. The implanted dose is high enough to override the



p-dopants and convert the SiGe source/drain of the NMOS into n-type 10' material, and the Si cap layer into highly n-doped 15'. The activation of this implant 58 is typically done with rapid thermal annealing, which allows for relatively precise control of the metallurgical junction location. In an exemplary embodiment rapid thermal anneal conditions can be about 1050-1100°C for less than 10sec. Or, one can use laser anneal or flash anneal to achieve even shorter annealing times.

There can be variations on the epi deposition and source/drain doping steps. If need arises instead of p-doped SiGe epitaxy, one can carry out undoped SiGe epitaxy and implant and activate the source/drain for both the PMOS and NMOS devices. For some situations one can even use n-type epi deposition. For one ordinarily skilled in the art further possible variations on such steps will be apparent.

Fig. 4 shows processing steps in the fabrication of an alternate exemplary embodiment of a heterojunction source/drain field device. Since as mentioned before the main beneficiary of the SiGe heterojunction source/drain is the PFET device, it is possible to omit the heterojunction from the NFET altogether, in order to avoid strain arising from the SiGe source/drain, as discussed in reference to Fig. 2B. Fig. 4A shows the devices at an intermediate state of processing, with all steps before the source/drain recession carried out as shown in Figs. 3, with the SiGe epi layer 20 and the Si epi layer 30 present in the NMOS. However the source/drain recessing and SiGe source/drain epi steps are performed only for the PMOS. This results in the situation shown in Fig. 4A, with the PMOS having the SiGe epi source/drain 10 in place, and the NMOS without

source and drain. The next step shown in Fig. 4B is similar to the one of Fig. 3E. The PMOS is masked 57, and the source/drain of the NFET is implanted 58. The difference between this embodiment and that of Fig.3E is that the implant 58 now simply goes into the Si Body 40', resulting in regular source/drain 70, with no heterojunction being formed with the Si body 40'.

Fig. 5 shows a symbolic view of a processor containing a chip which contains a heterojunction source/drain field device. Such a processor 900 has at least one chip 901, which contains at least one field effect device 100 having a SiGe heterojunction source/drain and with the coinciding heterojunction and metallurgical junction to within less than about 10nm, and preferably less than about 5nm of tolerance. The processor 900 can be any processor which can benefit from the SiGe heterojunction source/drain device 100. These devices form part of the processor in their multitude on one or more chips 901. In a representative embodiment these SiGe heterojunction source/drain field effect devices 100 are fabricated in an SOI technology. Representative embodiments of processors manufactured with SiGe heterojunction source/drain field effect devices are digital processors, typically found in the central processing complex of computers; mixed digital/analog processors, which benefit significantly from the high mobility of the carriers in p-type SiGe heterojunction source/drain field effect devices; and in general any communication processor, such as modules connecting memories to processors, routers, radar systems, high performance video-telephony, game modules, and others.

Many modifications and variations of the present invention are possible in light of

the above teachings, and could be apparent for those skilled in the art. The scope of the invention is defined by the appended claims.